RSAC 2025 Conference

Many Voices.
One Community.

SESSION ID: IAIS-T01

How Security UX Must Change, with Agentive AI

Steph Hay
Senior Director, Gemini Product & UX, Google Cloud Security

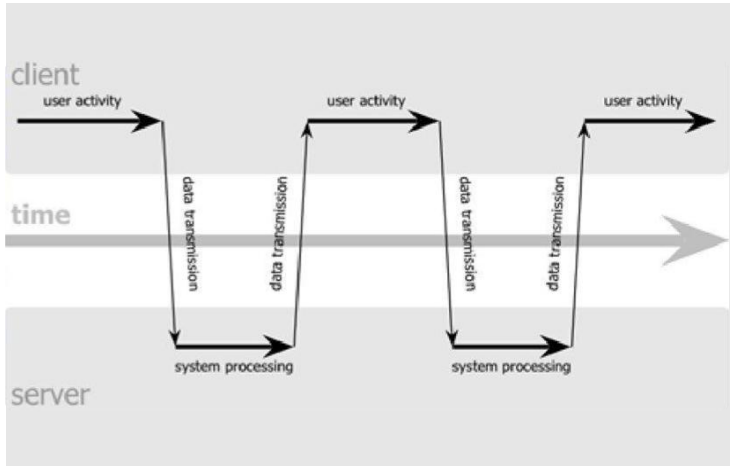
Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the presenters individually and, unless expressly stated to the contrary, are not the opinion or position of RSA Conference LLC or any other co-sponsors. RSA Conference LLC does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented.

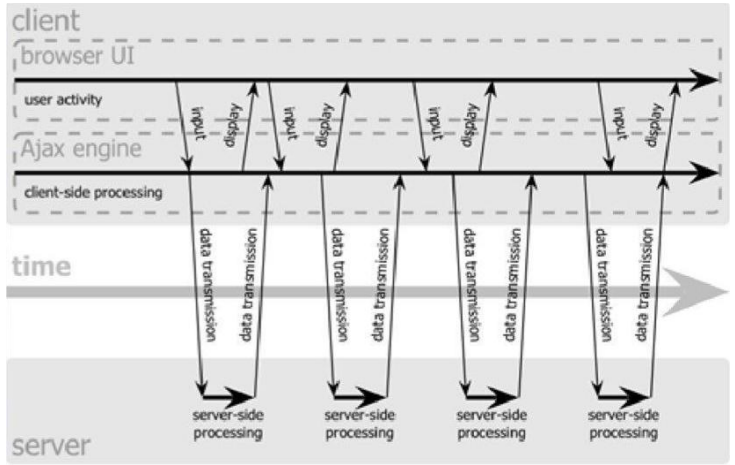
Attendees should note that sessions may be audio- or video-recorded and may be published in various media, including print, audio and video formats without further notice. The presentation template and any media capture are subject to copyright protection.

© 2025 RSA Conference LLC or its affiliates. The RSAC and RSAC CONFERENCE logos and other trademarks are proprietary. All rights reserved.

20 years ago dawned the era of the dynamic web



Classic web application model
(synchronous)

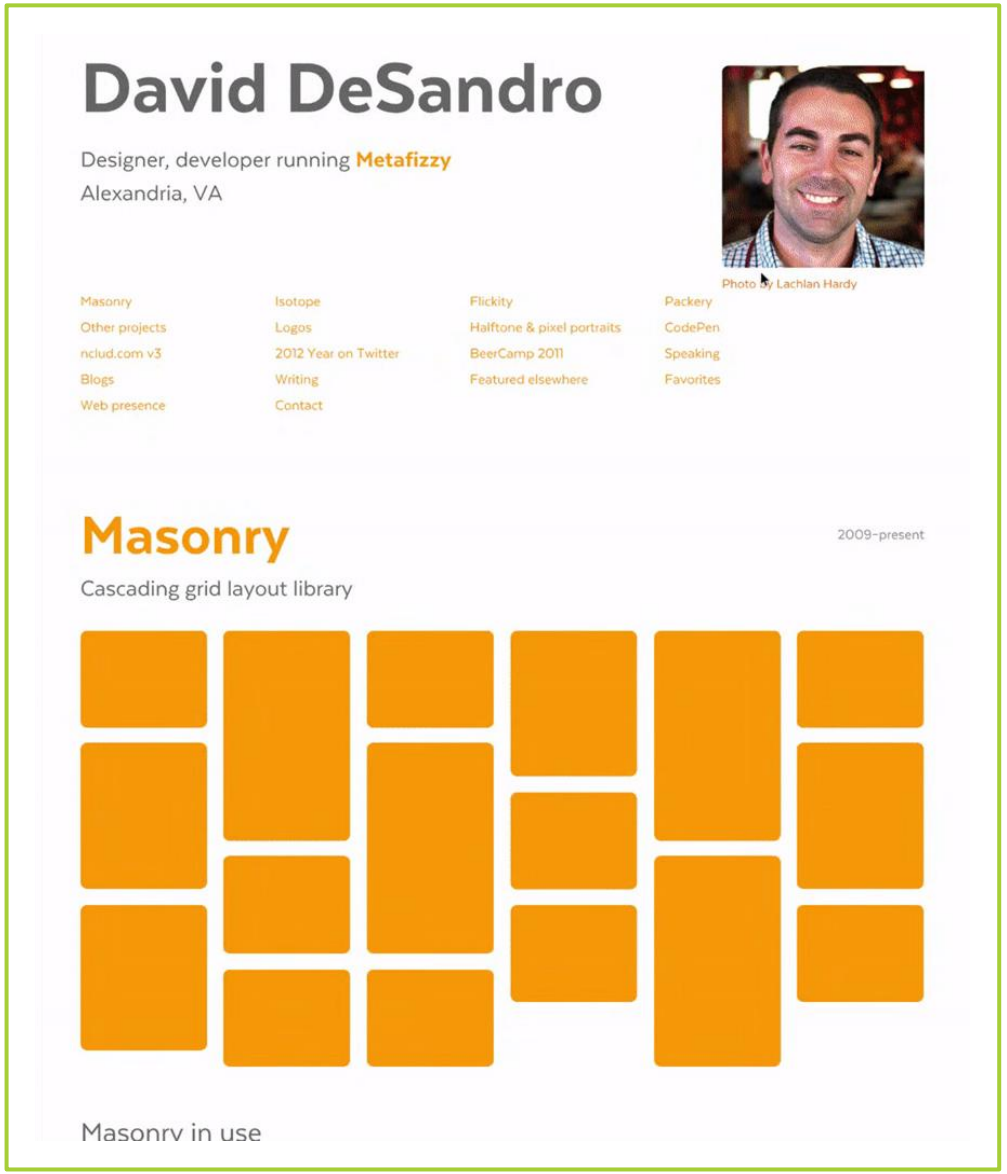


Ajax web application model
(asynchronous)

Source

Then came the mobile revolution…

- Thumbs navigate screens much differently than a mouse, thereby requiring new tactile and interaction modes
- The use cases switched from “sitting down to do work” to playing games and browsing “on the go”
- The focus became enabling features across different UIs (mobile, tablet, laptop, desktop) until new technologies enabled intuitive scaling



Thanks to generative AI, we're at a similar inflection point

Agentic micro UIs in action

Designing for Agentive UX — A Guide

AI experience? It's like saying electricity experience — too vague.

Kshitij Agrawal · Follow
Published in AI-UX Collective · 8 min read · Aug 18, 2024

Dialogflow > Documentation > ES

Was this helpful?

General agent design best practices

Send feedback

This guide provides general best practices for designing all types of agents.

You should also see the voice agent design guide specifically for designing voice agents, and the best practices guide for using the Dialogflow service.

The agentic era of UX

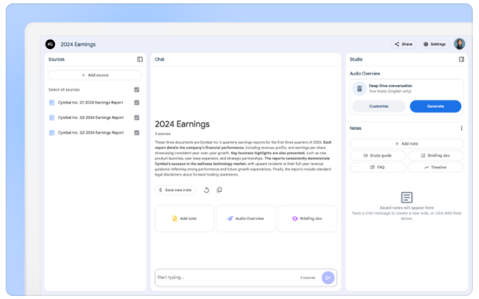
The future of digital experience is here — but it's being minced into microscopic use cases.

Alex Klein · Follow
Published in UX Collective · 8 min read · May 6, 2024

Unlock your organization's collective brilliance

New ways to interact and engage with your enterprise data using NotebookLM

We built NotebookLM to help users make sense of complex information, and now we're bringing this capability to enterprises. With NotebookLM Plus, your employees can upload information to synthesize, uncover insights, and enjoy new ways of engaging with data, such as podcast-like audio summaries and more. It's the same experience millions of users love, now enhanced with security and privacy features for work.



Agents represent a new era in human-computer interaction

Offloading Tasks

- Agents are goal-oriented, e.g., “Reduce time to detect” or “Triage and respond to low-risk alerts”
- Continuously learning and making decisions based on context, which can and does change

Dynamic UIs

- Infused in the existing practitioners’ tooling, both proactively guiding and invisibly advancing workflows
- Foreground their decisions as a means to develop understanding and boost confidence

Exponential Outcomes

- Learning and generative responses outpace traditional software development processes
- Users are more productive without having to learn AI
- Organizations are more [fortified] without having to hire more staff

And not a moment too soon…

Threats

Growing number of threats, increasingly sophisticated

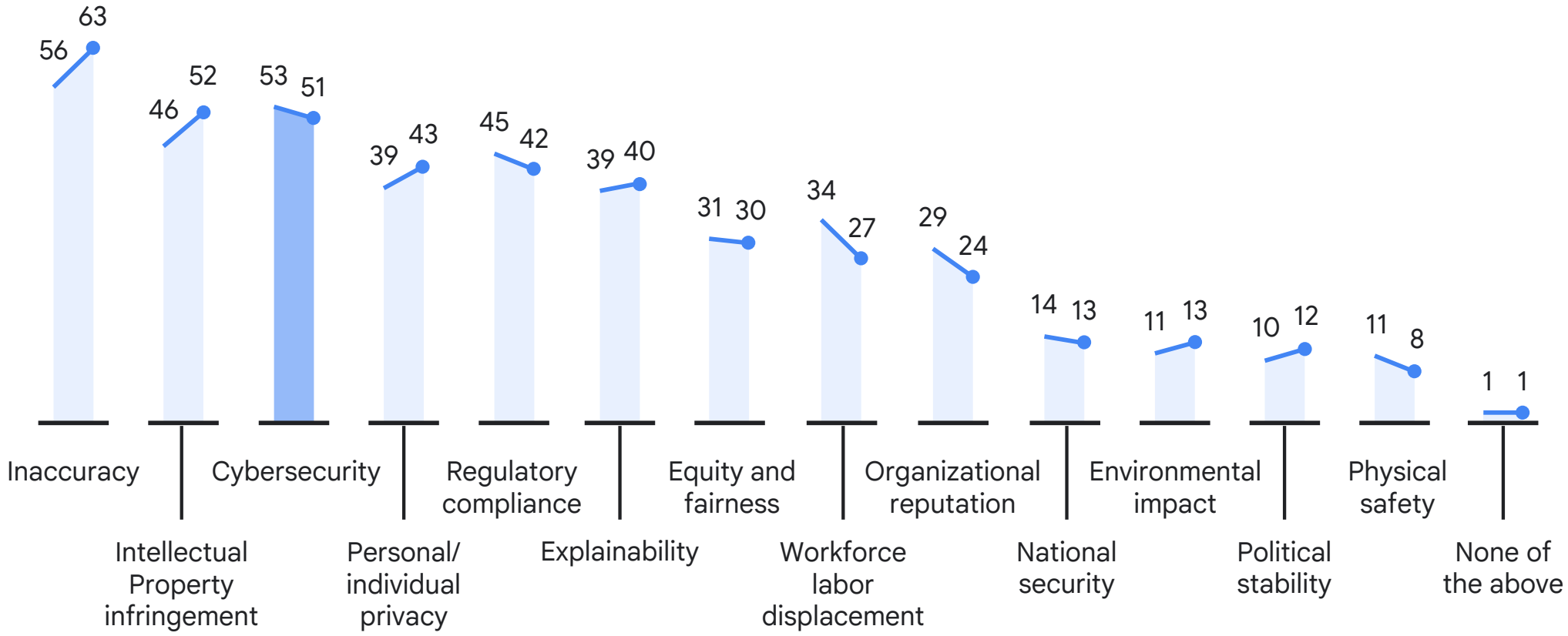
Toil

Burdensome, repetitive tasks and manual steps to resolve issues

Talent

Lack of experts, high turnover

Security, privacy and compliance are ranked among the biggest Gen AI risks

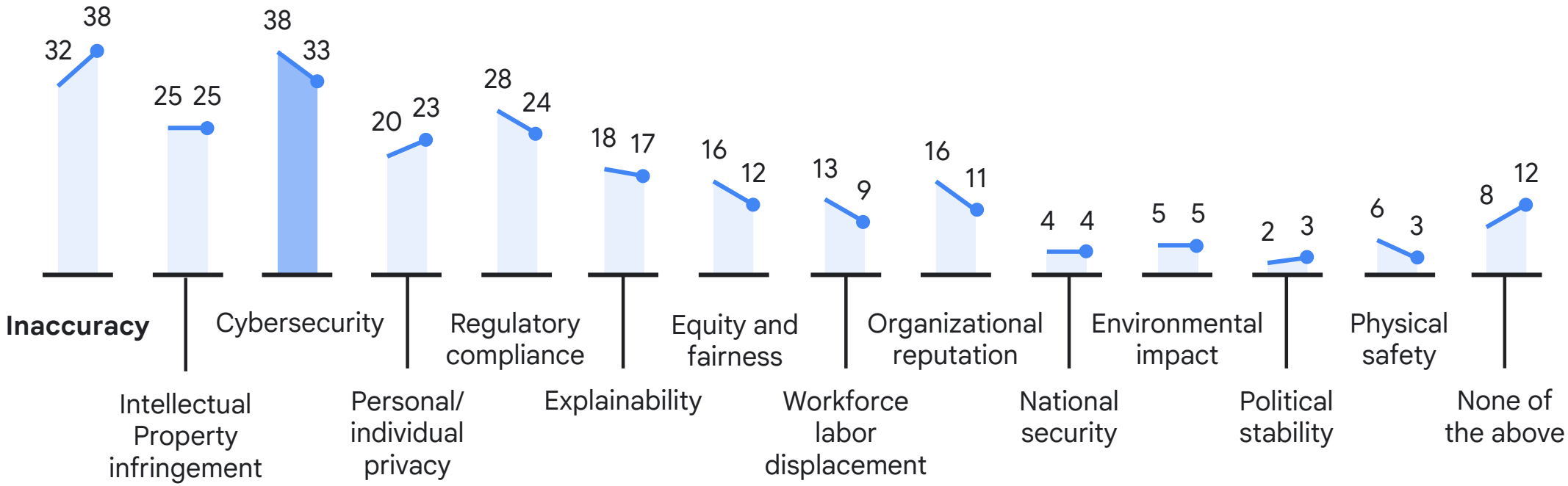


Gen AI risks that organizations consider relevant, [1] % respondents

2023 — 2024

Source: McKinsey Global Survey on AI
1,363 participants at all levels of the organization, Feb 22-Mar5, 2024

Security is cited as the 2nd highest risk customers are actively working to mitigate



Gen AI risks that organizations are working to mitigate, [1] % respondents

2023 — 2024

Source: McKinsey Global Survey on AI
1,363 participants at all levels of the organization, Feb 22-Mar5, 2024

Shifting from manual to semi-autonomous

Manual

Humans cyclically and routinely execute tasks

Assisted

Humans still do their typical job, but AI boosts their productivity by generating summaries, recommendations, etc.

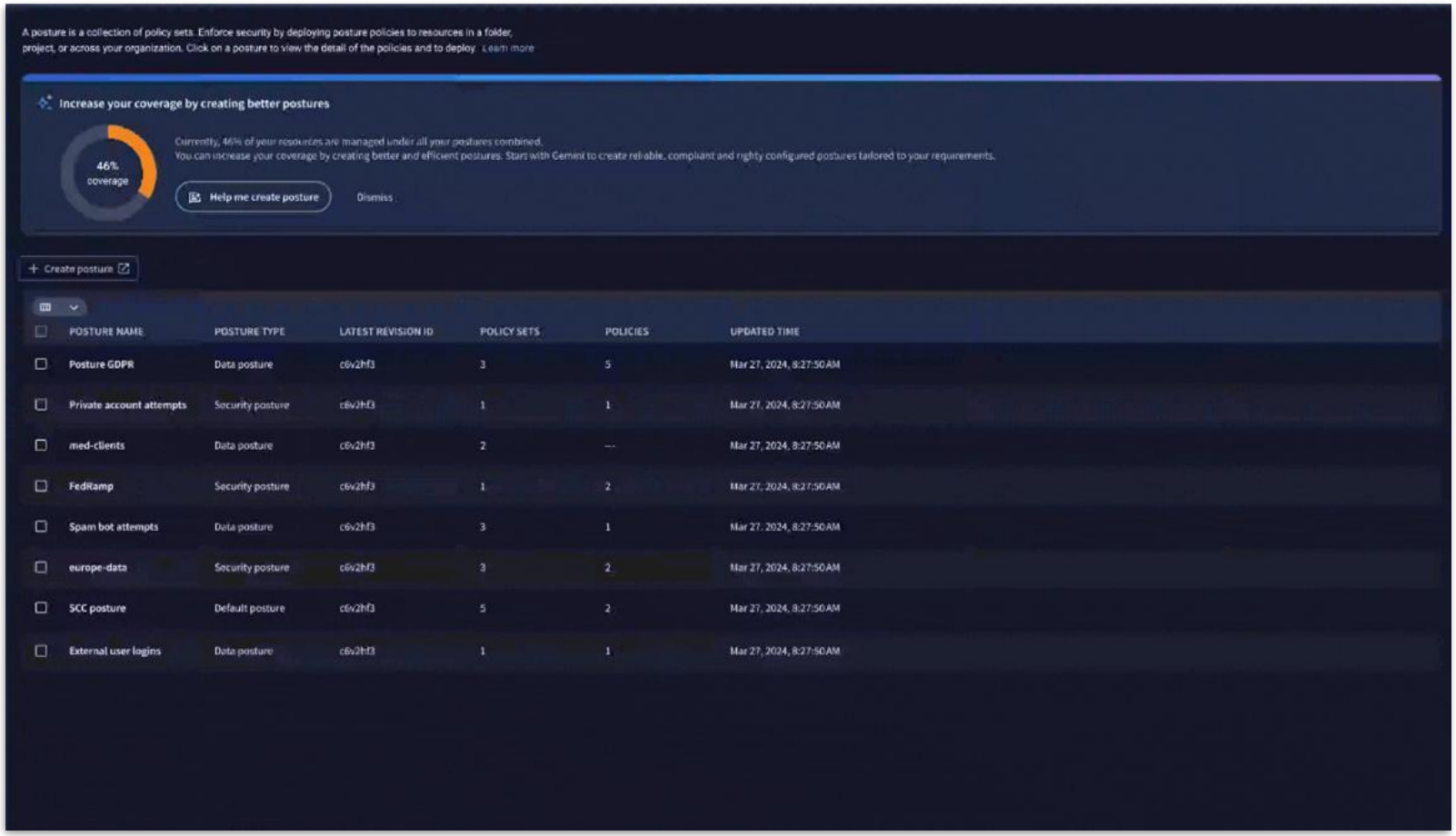
Semi- autonomous

AI drives most tasks across the security lifecycle consistently well, delegating tasks it can't automate with sufficiently high confidence or precision

Autonomous

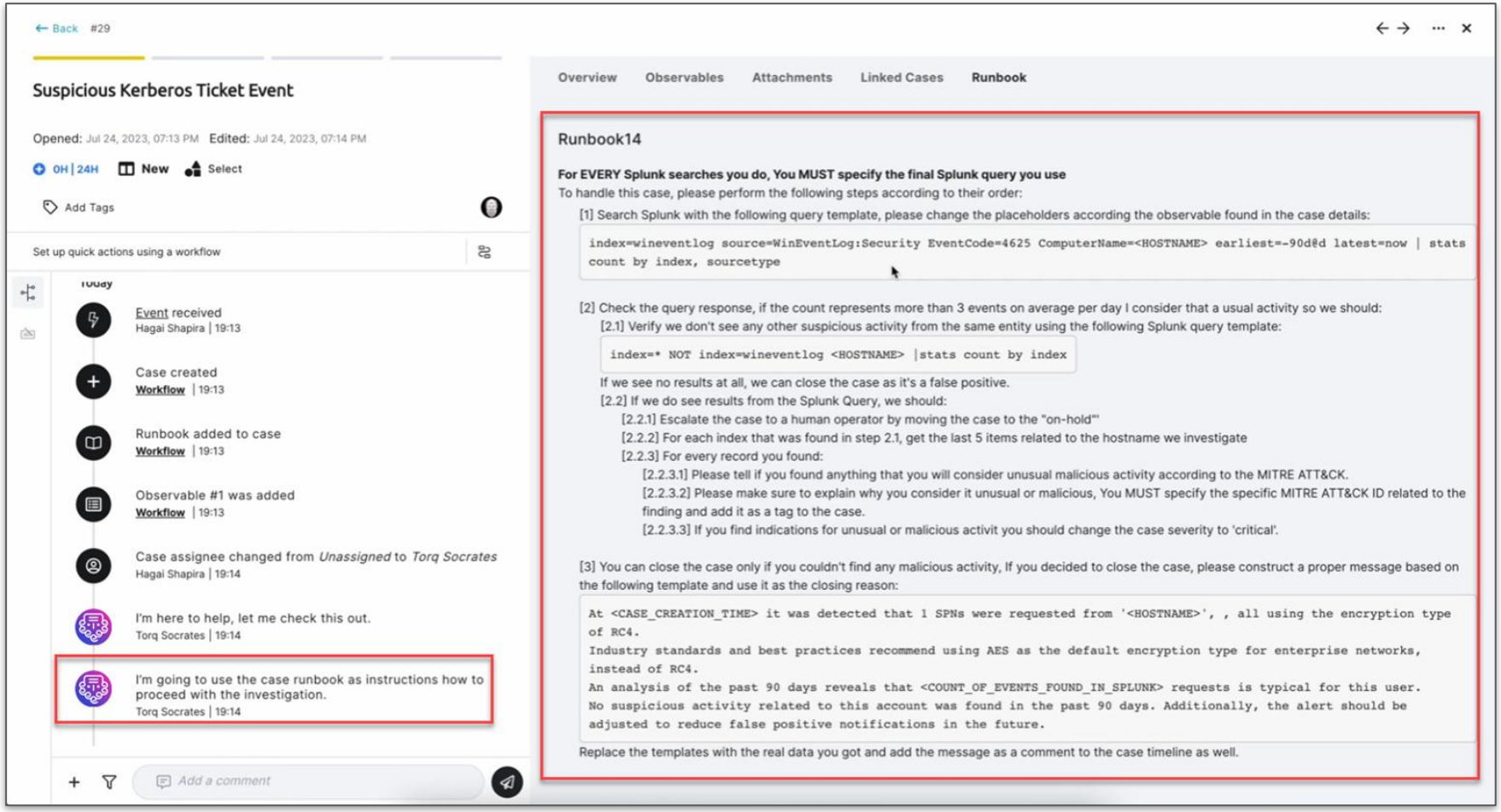
AI drives the security lifecycle to positive outcomes on behalf of users

Assistive UX is security is necessary!



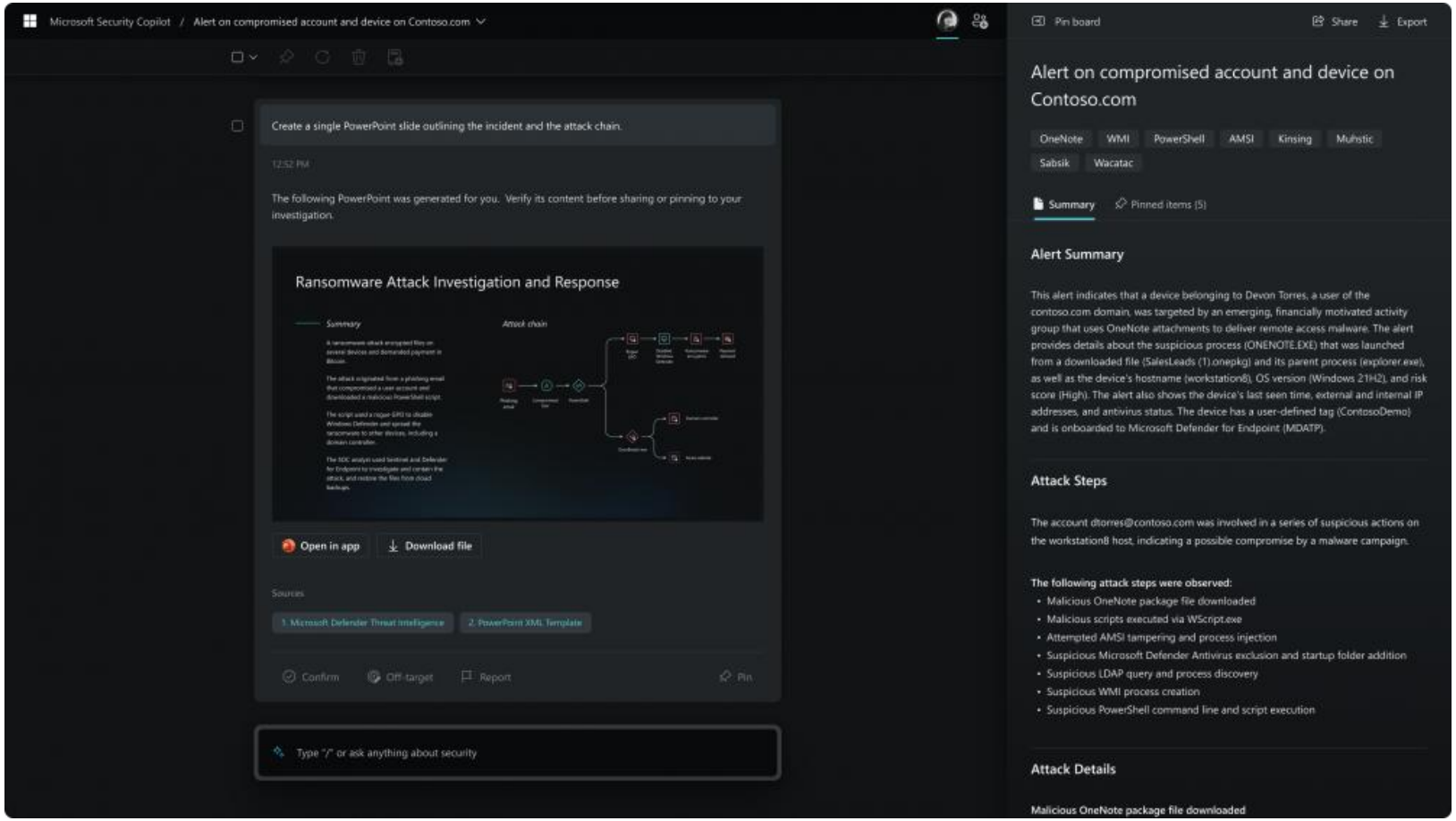
"easy button" infused in existing workflows

Assistive UX is security is necessary!


Back #29
Suspicious Kerberos Ticket Event
Opened: Jul 24, 2023, 07:13 PM Edited: Jul 24, 2023, 07:14 PM
0H | 24H
New
Select
Add Tags
Set up quick actions using a workflow
Today
Event received
Hagai Shapira | 19:13
Case created
Workflow | 19:13
Runbook added to case
Workflow | 19:13
Observable #1 was added
Workflow | 19:13
Case assignee changed from Unassigned to Torq Socrates
Hagai Shapira | 19:14
I'm here to help, let me check this out.
Torq Socrates | 19:14
I'm going to use the case runbook as instructions how to proceed with the investigation.
Torq Socrates | 19:14
Add a comment
Overview
Observables
Attachments
Linked Cases
Runbook
Runbook14
For EVERY Splunk searches you do, You MUST specify the final Splunk query you use
To handle this case, please perform the following steps according to their order:
[1] Search Splunk with the following query template, please change the placeholders according the observable found in the case details:
index=wineventlog source=WinEventLog:Security EventCode=4625 ComputerName=<HOSTNAME> earliest=-90d@d latest=now | stats count by index, sourcetype
[2] Check the query response, if the count represents more than 3 events on average per day I consider that a usual activity so we should:
[2.1] Verify we don't see any other suspicious activity from the same entity using the following Splunk query template:
index=* NOT index=wineventlog <HOSTNAME> |stats count by index
If we see no results at all, we can close the case as it's a false positive.
[2.2] If we do see results from the Splunk Query, we should:
[2.2.1] Escalate the case to a human operator by moving the case to the "on-hold""
[2.2.2] For each index that was found in step 2.1, get the last 5 items related to the hostname we investigate
[2.2.3] For every record you found:
[2.2.3.1] Please tell if you found anything that you will consider unusual malicious activity according to the MITRE ATT&CK.
[2.2.3.2] Please make sure to explain why you consider it unusual or malicious, You MUST specify the specific MITRE ATT&CK ID related to the finding and add it as a tag to the case.
[2.2.3.3] If you find indications for unusual or malicious activit you should change the case severity to 'critical'.
[3] You can close the case only if you couldn't find any malicious activity, If you decided to close the case, please construct a proper message based on the following template and use it as the closing reason:
At <CASE_CREATION_TIME> it was detected that 1 SPNs were requested from '<HOSTNAME>', , all using the encryption type of RC4.
Industry standards and best practices recommend using AES as the default encryption type for enterprise networks, instead of RC4.
An analysis of the past 90 days reveals that <COUNT_OF_EVENTS_FOUND_IN_SPLUNK> requests is typical for this user.
No suspicious activity related to this account was found in the past 90 days. Additionally, the alert should be adjusted to reduce false positive notifications in the future.
Replace the templates with the real data you got and add the message as a comment to the case timeline as well.

seeded prompts to enable faster adoption given state of quality

Assistive UX is security is necessary!


Microsoft Security Copilot / Alert on compromised account and device on Contoso.com
Create a single PowerPoint slide outlining the incident and the attack chain.
The following PowerPoint was generated for you. Verify its content before sharing or pinning to your investigation.
Ransomware Attack Investigation and Response
Summary
Attack chain
Open in app
Download file
Sources
Confirm
Off-target
Report
Pin
Type "/" or ask anything about security
Pin board
Share
Export
Alert on compromised account and device on Contoso.com
OneNote
WMI
PowerShell
AMSI
Kinsing
Mubstic
Sabsik
Wacatac
Summary
Pinned items (5)
Alert Summary
This alert indicates that a device belonging to Devon Torres, a user of the contoso.com domain, was targeted by an emerging, financially motivated activity group that uses OneNote attachments to deliver remote access malware. The alert provides details about the suspicious process (ONENOTE.EXE) that was launched from a downloaded file (SalesLeads (1).onepkg) and its parent process (explorer.exe), as well as the device's hostname (workstation8), OS version (Windows 21H2), and risk score (High). The alert also shows the device's last seen time, external and internal IP addresses, and antivirus status. The device has a user-defined tag (ContosoDemo) and is onboarded to Microsoft Defender for Endpoint (MDATP).
Attack Steps
The account dtorres@contoso.com was involved in a series of suspicious actions on the workstation8 host, indicating a possible compromise by a malware campaign.
The following attack steps were observed:
Malicious OneNote package file downloaded
Malicious scripts executed via WScript.exe
Attempted AMSI tampering and process injection
Suspicious Microsoft Defender Antivirus exclusion and startup folder addition
Suspicious LDAP query and process discovery
Suspicious WMI process creation
Suspicious PowerShell command line and script execution
Attack Details
Malicious OneNote package file downloaded

BP
Write a YARA-L rule to detect potential Cobalt Strike beaconing activity based on suspicious network connections to uncommon ports and specific user-agent strings in network logs.
Enter a prompt here

multi-turn chat–both standalone and infused in the tooling

Unanticipated, positive outcomes from assistive AI

Agents are here to change security for good

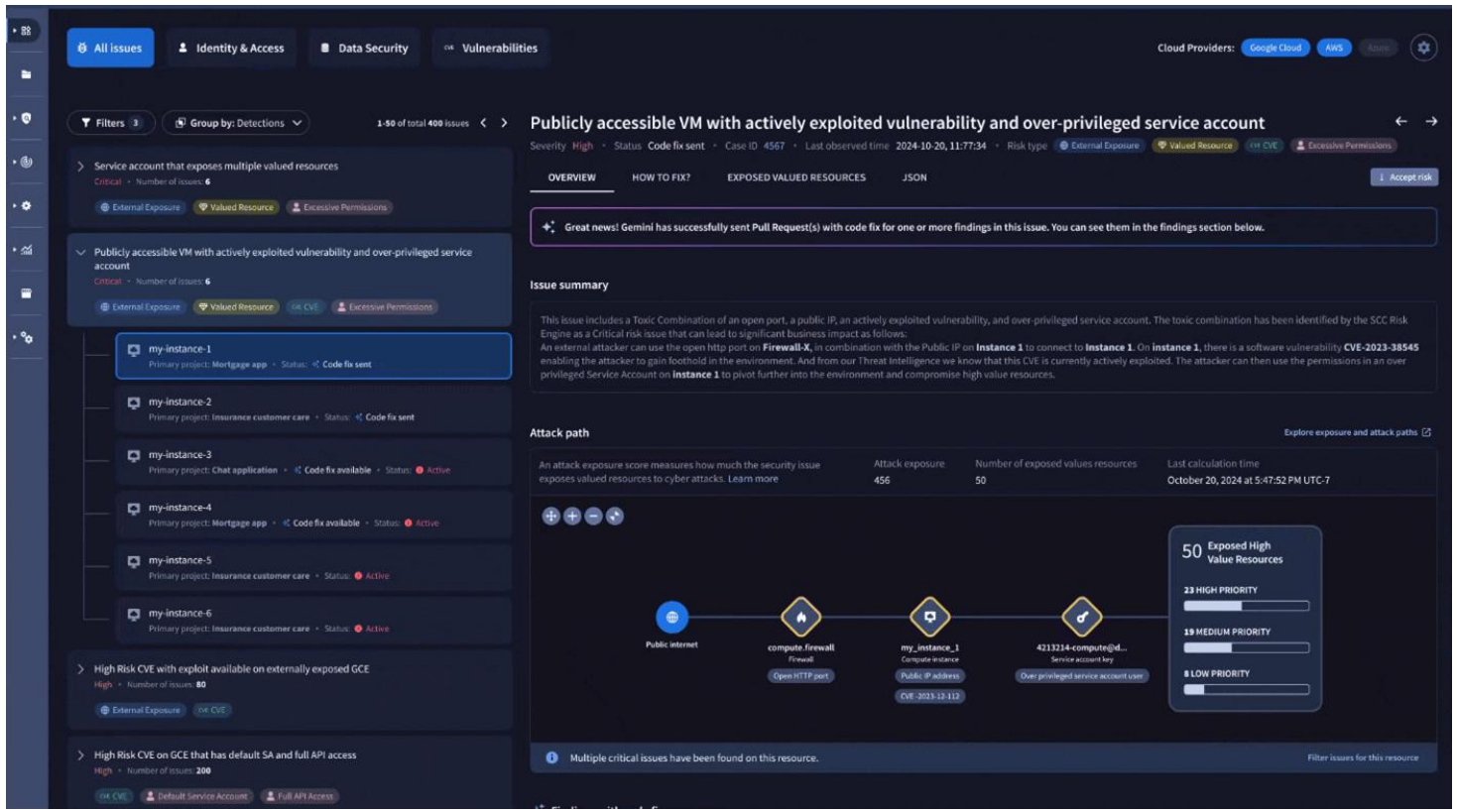
NEVER DRIFT FROM YOUR IDEAL STATE. Only update org and access policies as business needs change because AI manages least-privilege for you.

NEVER WRITE ANOTHER RULE. Only write or tune detections for your unique use cases because AI writes and updates rules for you.

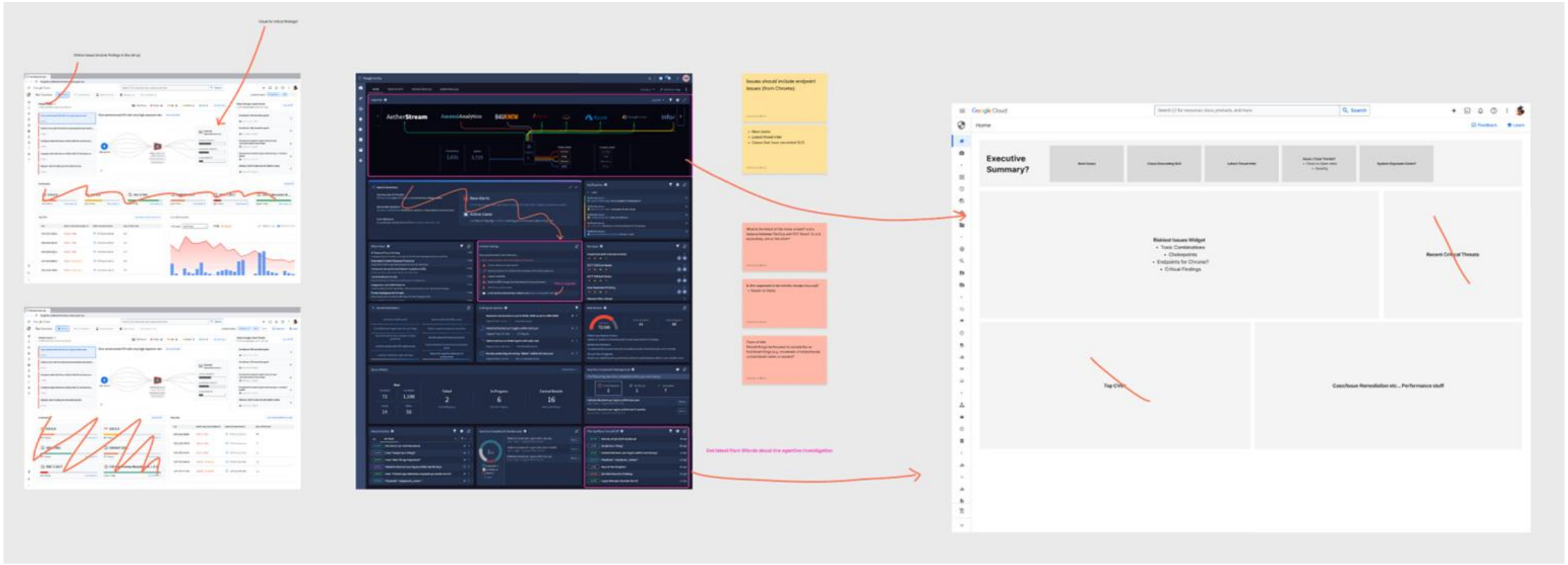
NEVER INVESTIGATE NOISE. Only see high-value, de-duplicated issues because AI always identifies the right risks relevant to your business.

NEVER WRITE ANOTHER PLAYBOOK. Only spend time containing incidents because AI creates/runs playbooks, updates controls, and generates audit reports.

Showing up in security tools to offload low-value tasks from practitioners…

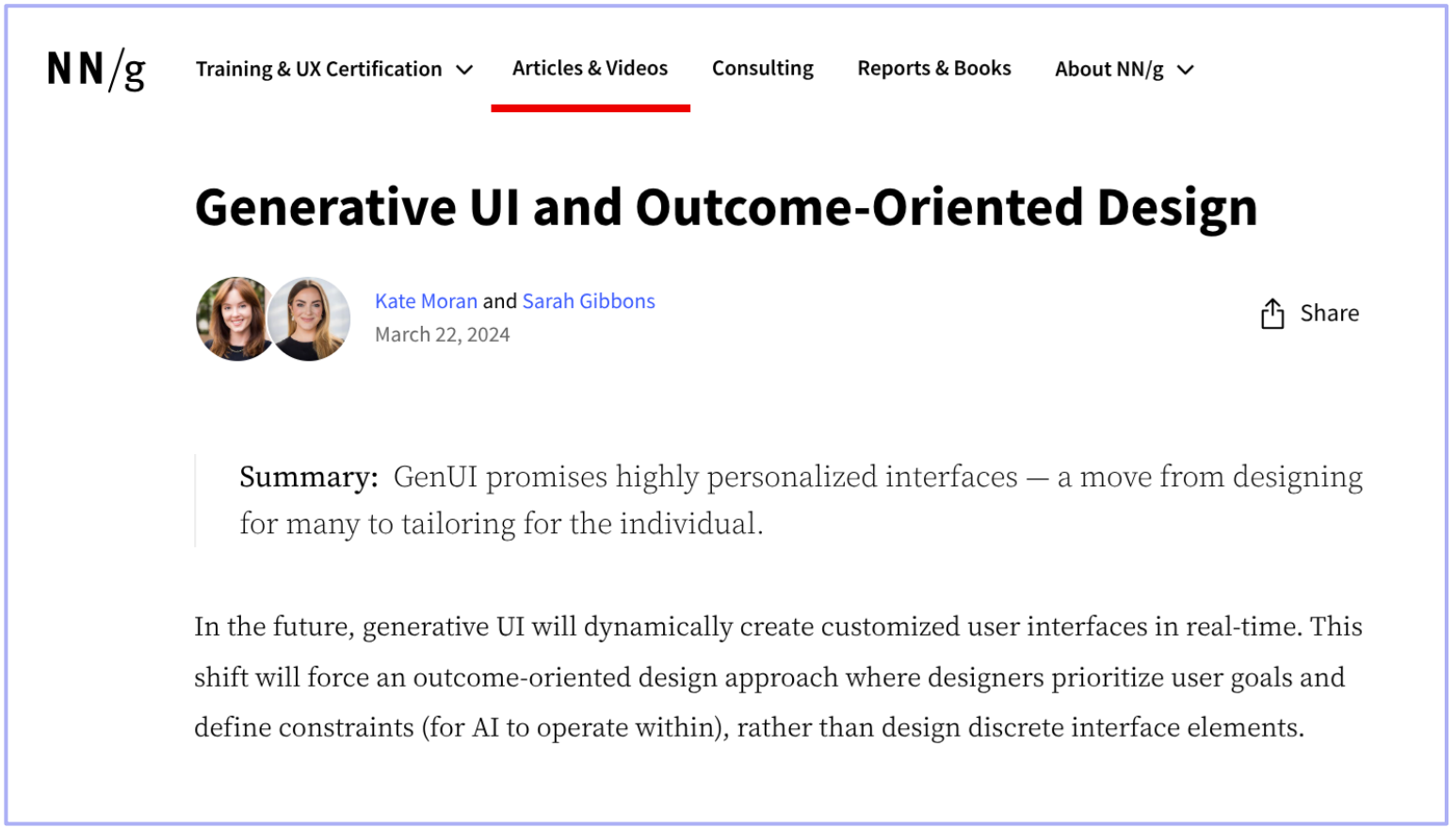


However! Our UIs aren't designed to be multi-turn or generative by default…



So our UIs must be adaptive to dynamic content that changes over time

- Any UI is a dynamic container to house generative content relevant to the individual
- That content will include a combination of agent-driven recommendations and information, in response to the individual human users inputs and behaviors
- Agents will change HOW they render outputs over time



NN/g Training & UX Certification Articles & Videos Consulting Reports & Books About NN/g

Generative UI and Outcome-Oriented Design

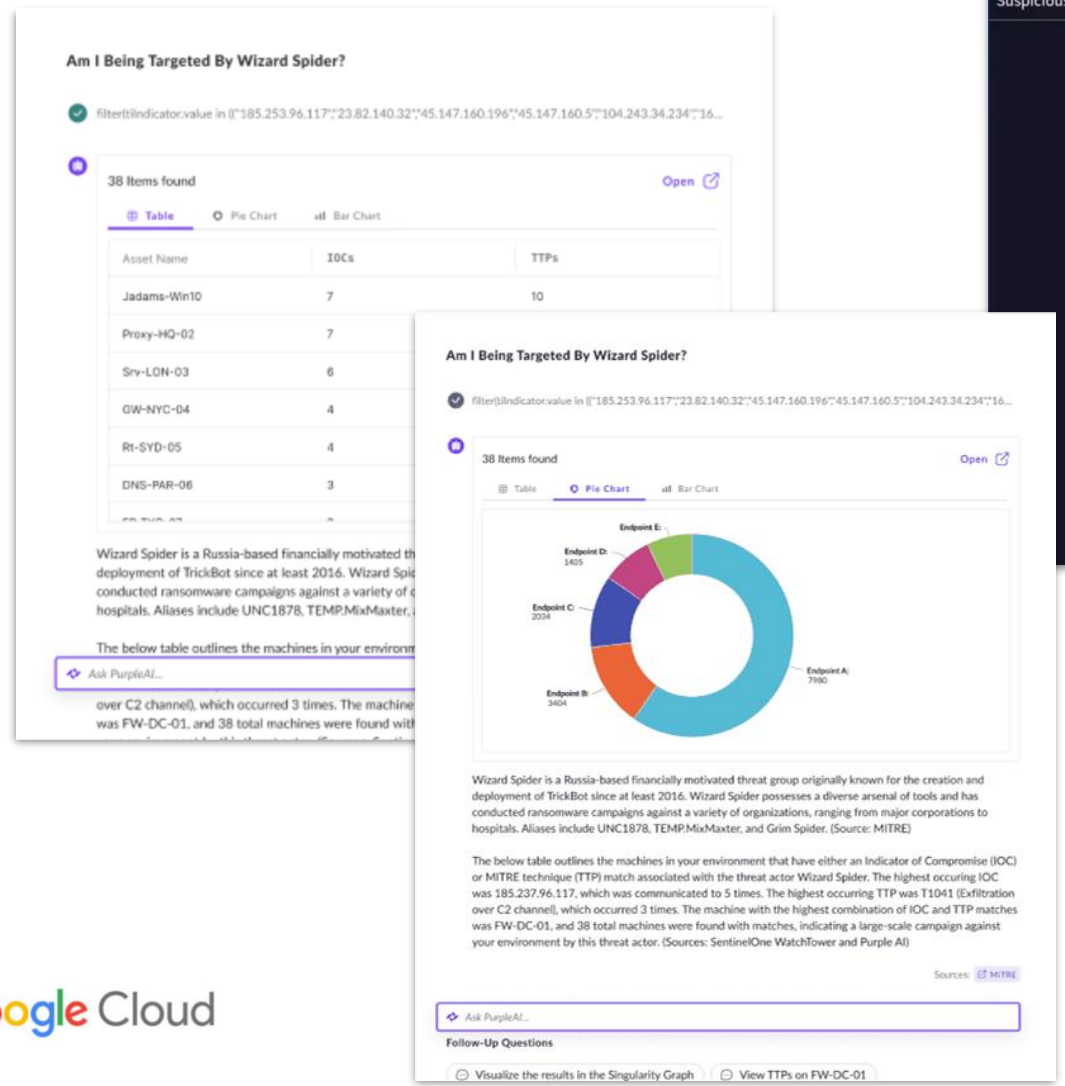
Kate Moran and Sarah Gibbons
March 22, 2024

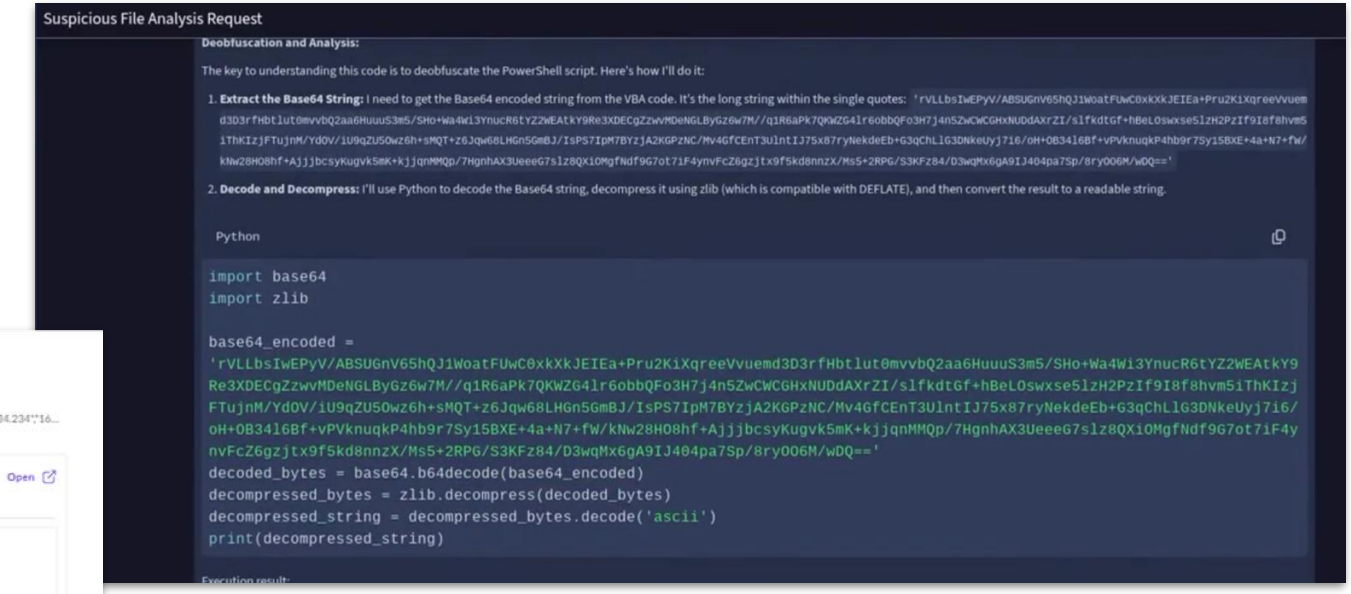
Share

Summary: GenUI promises highly personalized interfaces — a move from designing for many to tailoring for the individual.

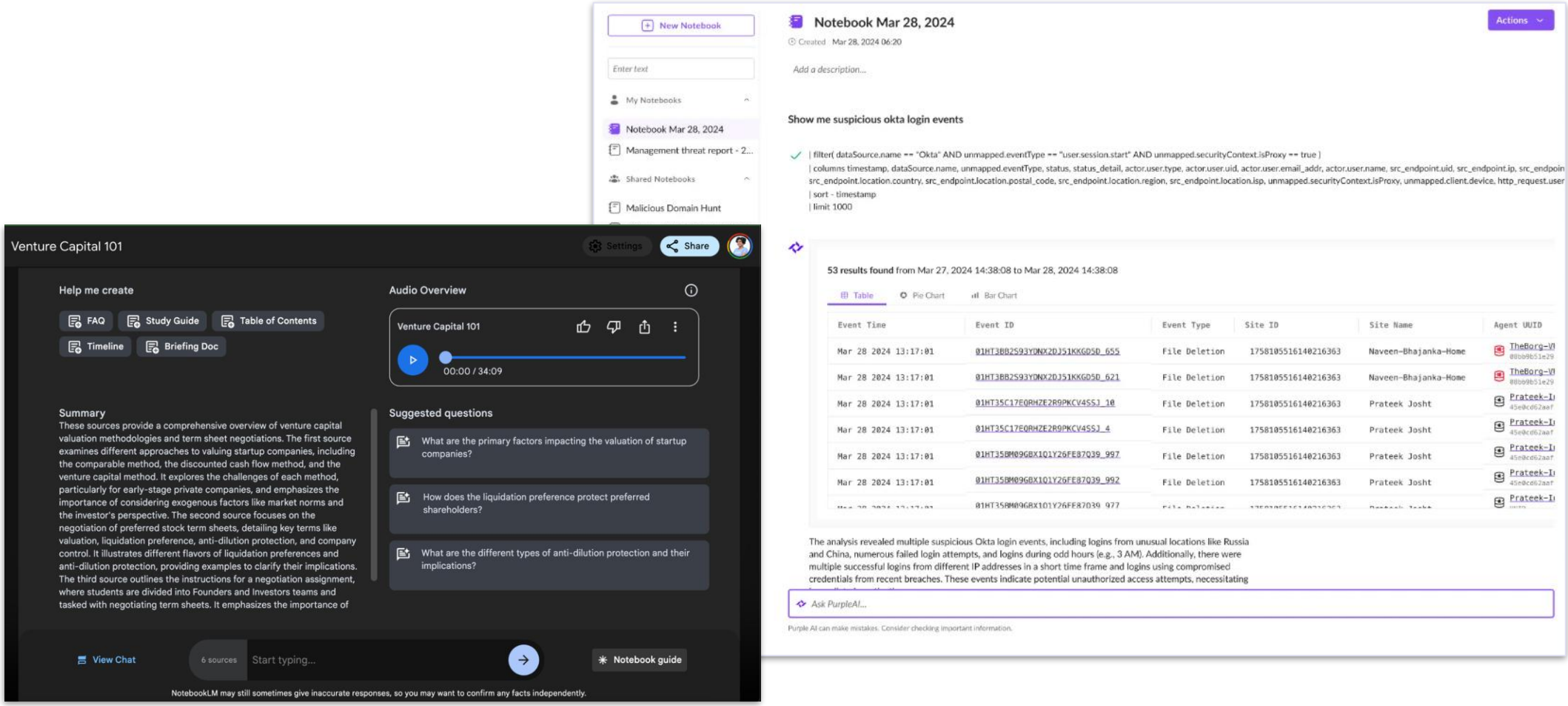
In the future, generative UI will dynamically create customized user interfaces in real-time. This shift will force an outcome-oriented design approach where designers prioritize user goals and define constraints (for AI to operate within), rather than design discrete interface elements.

Early examples of new interaction models: multi-modal content


Am I Being Targeted By Wizard Spider?
38 Items found
Open
Table
Pie Chart
Bar Chart
Asset Name
IOCs
TTPs
Jadams-Win10
7
10
Proxy-HQ-02
7
Srv-LON-03
6
GW-NYC-04
4
Rt-SYD-05
4
DNS-PAR-06
3
Endpoint E:
Endpoint D:
1405
Endpoint C:
2034
Endpoint B:
3404
Endpoint A:
7980
Wizard Spider is a Russia-based financially motivated threat group originally known for the creation and deployment of TrickBot since at least 2016. Wizard Spider possesses a diverse arsenal of tools and has conducted ransomware campaigns against a variety of organizations, ranging from major corporations to hospitals. Aliases include UNC1878, TEMP.MixMaxter, and Grim Spider. (Source: MITRE)
The below table outlines the machines in your environment that have either an Indicator of Compromise (IOC) or MITRE technique (TTP) match associated with the threat actor Wizard Spider. The highest occuring IOC was 185.237.96.117, which was communicated to 5 times. The highest occurring TTP was T1041 (Exfiltration over C2 channel), which occurred 3 times. The machine with the highest combination of IOC and TTP matches was FW-DC-01, and 38 total machines were found with matches, indicating a large-scale campaign against your environment by this threat actor. (Sources: SentinelOne WatchTower and Purple AI)
Sources:
MITRE
Ask PurpleAI...
Follow-Up Questions
Visualize the results in the Singularity Graph
View TTPs on FW-DC-01

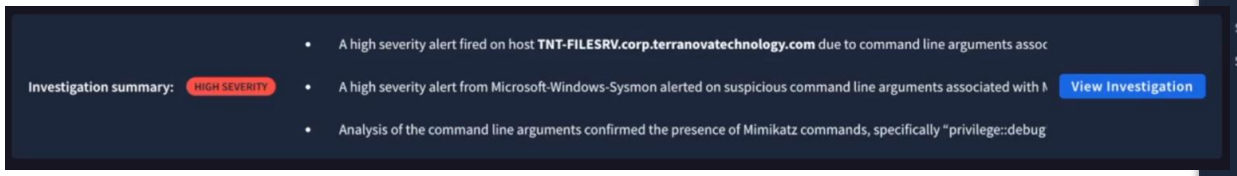


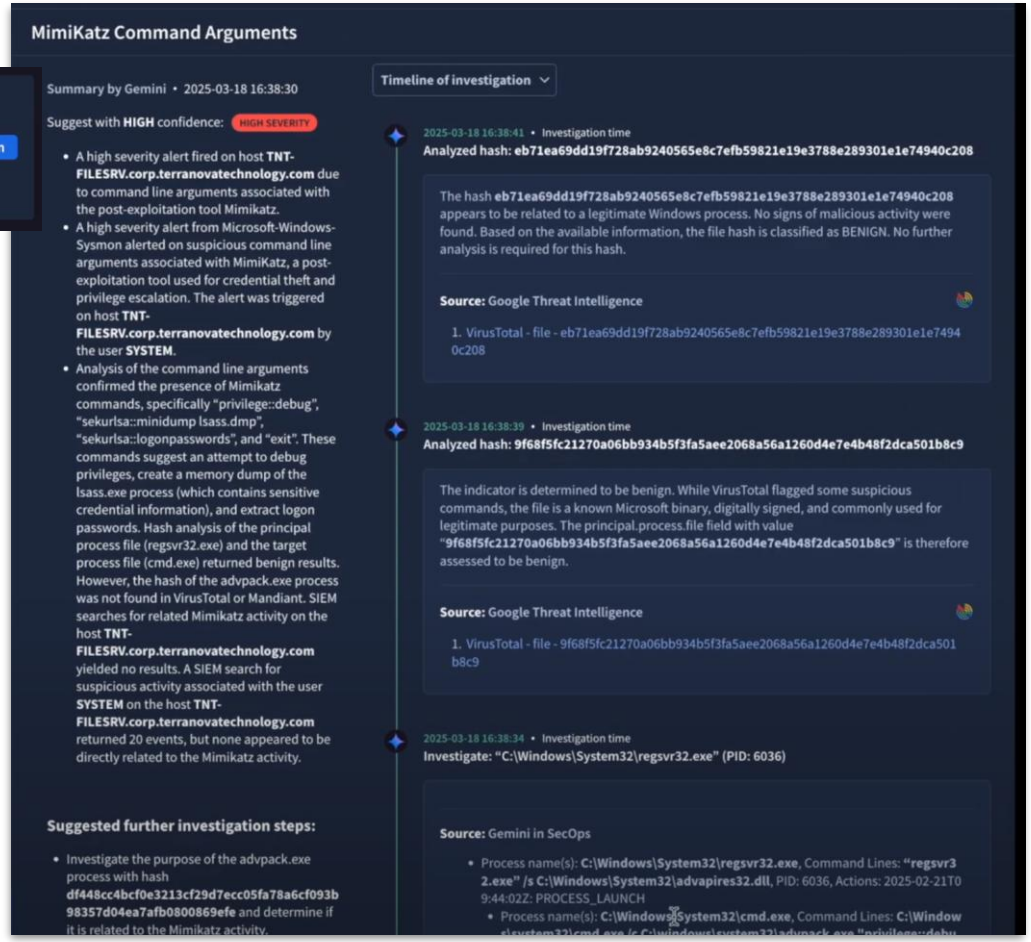
Early examples of new interaction models: collaborative, multi-modal notebooks


New Notebook
Enter text
My Notebooks
Notebook Mar 28, 2024
Management threat report - 2...
Shared Notebooks
Malicious Domain Hunt
Notebook Mar 28, 2024
Actions
Created Mar 28, 2024 06:20
Add a description...
Show me suspicious okta login events
| filter(dataSource.name == "Okta" AND unmapped.eventType == "user.session.start" AND unmapped.securityContext.isProxy == true)
| columns timestamp, dataSource.name, unmapped.eventType, status, status_detail, actor.user.type, actor.user.uid, actor.user.email_addr, actor.user.name, src_endpoint.uid, src_endpoint.ip, src_endpoint.location.country, src_endpoint.location.postal_code, src_endpoint.location.region, src_endpoint.location.isp, unmapped.securityContext.isProxy, unmapped.client.device, http_request.user
| sort - timestamp
| limit 1000
53 results found from Mar 27, 2024 14:38:08 to Mar 28, 2024 14:38:08
Table
Pie Chart
Bar Chart
Event Time
Event ID
Event Type
Site ID
Site Name
Agent UUID
Mar 28 2024 13:17:01
File Deletion
1758105516140216363
Naveen-Bhajanka-Home
Prateek Josht
The analysis revealed multiple suspicious Okta login events, including logins from unusual locations like Russia and China, numerous failed login attempts, and logins during odd hours (e.g., 3 AM). Additionally, there were multiple successful logins from different IP addresses in a short time frame and logins using compromised credentials from recent breaches. These events indicate potential unauthorized access attempts, necessitating
Ask PurpleAI...
Purple AI can make mistakes. Consider checking important information.
Venture Capital 101
Settings
Share
Help me create
FAQ
Study Guide
Table of Contents
Timeline
Briefing Doc
Audio Overview
Venture Capital 101
00:00 / 34:09
Summary
These sources provide a comprehensive overview of venture capital valuation methodologies and term sheet negotiations. The first source examines different approaches to valuing startup companies, including the comparable method, the discounted cash flow method, and the venture capital method. It explores the challenges of each method, particularly for early-stage private companies, and emphasizes the importance of considering exogenous factors like market norms and the investor's perspective. The second source focuses on the negotiation of preferred stock term sheets, detailing key terms like valuation, liquidation preference, anti-dilution protection, and company control. It illustrates different flavors of liquidation preferences and anti-dilution protection, providing examples to clarify their implications. The third source outlines the instructions for a negotiation assignment, where students are divided into Founders and Investors teams and tasked with negotiating term sheets. It emphasizes the importance of
Suggested questions
What are the primary factors impacting the valuation of startup companies?
How does the liquidation preference protect preferred shareholders?
What are the different types of anti-dilution protection and their implications?
View Chat
6 sources
Start typing...
Notebook guide
NotebookLM may still sometimes give inaccurate responses, so you may want to confirm any facts independently.

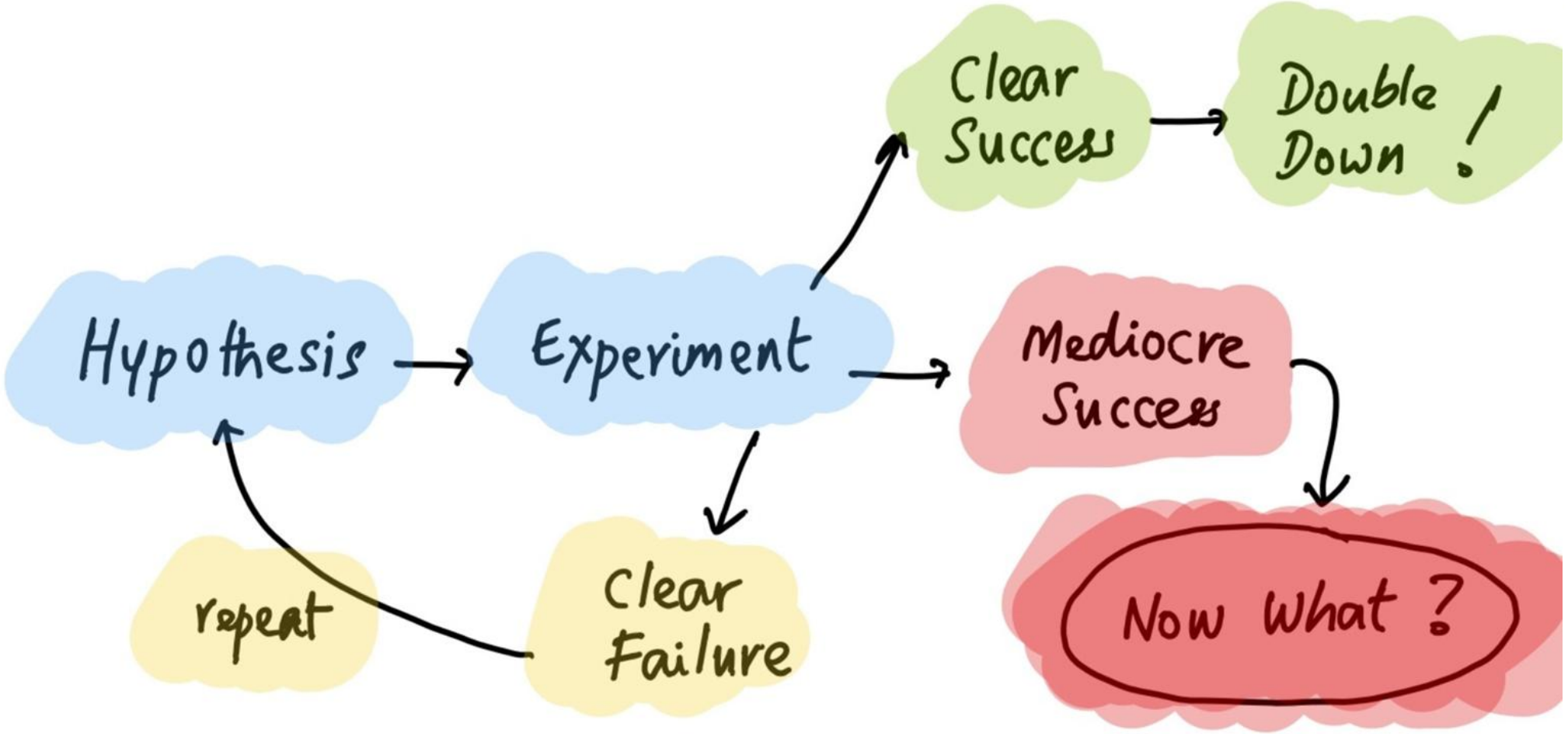
source

Early examples of new interaction models: investigation histories


Investigation summary:
HIGH SEVERITY
A high severity alert fired on host TNT-FILESRV.corp.terranovatechnology.com due to command line arguments assoc
A high severity alert from Microsoft-Windows-Sysmon alerted on suspicious command line arguments associated with M
View Investigation
Analysis of the command line arguments confirmed the presence of Mimikatz commands, specifically "privilege::debug


MimiKatz Command Arguments
Summary by Gemini • 2025-03-18 16:38:30
Suggest with HIGH confidence:
HIGH SEVERITY
A high severity alert fired on host TNT-FILESRV.corp.terranovatechnology.com due to command line arguments associated with the post-exploitation tool Mimikatz.
A high severity alert from Microsoft-Windows-Sysmon alerted on suspicious command line arguments associated with MimiKatz, a post-exploitation tool used for credential theft and privilege escalation. The alert was triggered on host TNT-FILESRV.corp.terranovatechnology.com by the user SYSTEM.
Analysis of the command line arguments confirmed the presence of Mimikatz commands, specifically "privilege::debug", "sekurlsa::minidump lsass.dmp", "sekurlsa::logonpasswords", and "exit". These commands suggest an attempt to debug privileges, create a memory dump of the lsass.exe process (which contains sensitive credential information), and extract logon passwords. Hash analysis of the principal process file (regsvr32.exe) and the target process file (cmd.exe) returned benign results. However, the hash of the advpack.exe process was not found in VirusTotal or Mandiant. SIEM searches for related Mimikatz activity on the host TNT-FILESRV.corp.terranovatechnology.com yielded no results. A SIEM search for suspicious activity associated with the user SYSTEM on the host TNT-FILESRV.corp.terranovatechnology.com returned 20 events, but none appeared to be directly related to the Mimikatz activity.
Suggested further investigation steps:
Investigate the purpose of the advpack.exe process with hash df448cc4bcf0e3213cf29d7ecc05fa78a6cf093b98357d04ea7afb0800869efe and determine if it is related to the Mimikatz activity.
Timeline of investigation
2025-03-18 16:38:41 • Investigation time
Analyzed hash: eb71ea69dd19f728ab9240565e8c7efb59821e19e3788e289301e1e74940c208
The hash eb71ea69dd19f728ab9240565e8c7efb59821e19e3788e289301e1e74940c208 appears to be related to a legitimate Windows process. No signs of malicious activity were found. Based on the available information, the file hash is classified as BENIGN. No further analysis is required for this hash.
Source: Google Threat Intelligence
1. VirusTotal - file - eb71ea69dd19f728ab9240565e8c7efb59821e19e3788e289301e1e74940c208
2025-03-18 16:38:39 • Investigation time
Analyzed hash: 9f68f5fc21270a06bb934b5f3fa5aee2068a56a1260d4e7e4b48f2dca501b8c9
The indicator is determined to be benign. While VirusTotal flagged some suspicious commands, the file is a known Microsoft binary, digitally signed, and commonly used for legitimate purposes. The principal.process.file field with value "9f68f5fc21270a06bb934b5f3fa5aee2068a56a1260d4e7e4b48f2dca501b8c9" is therefore assessed to be benign.
Source: Google Threat Intelligence
1. VirusTotal - file - 9f68f5fc21270a06bb934b5f3fa5aee2068a56a1260d4e7e4b48f2dca501b8c9
2025-03-18 16:38:34 • Investigation time
Investigate: "C:\Windows\System32\regsvr32.exe" (PID: 6036)
Source: Gemini in SecOps
Process name(s): C:\Windows\System32\regsvr32.exe, Command Lines: "regsvr32.exe" /s C:\Windows\System32\advapires32.dll, PID: 6036, Actions: 2025-02-21T09:44:02Z: PROCESS_LAUNCH
Process name(s): C:\Windows\System32\cmd.exe, Command Lines: C:\Windows

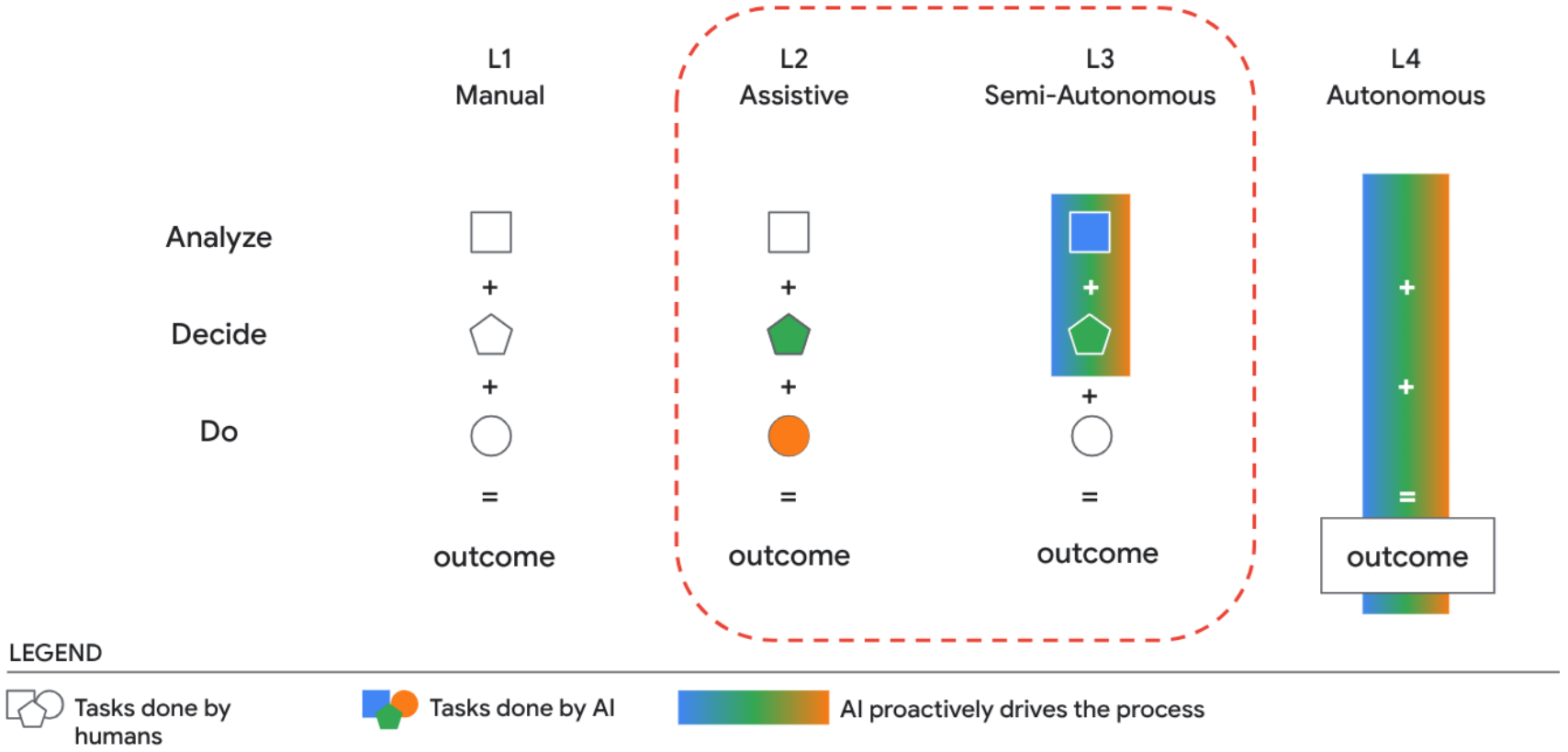
What's coming next?



source

Measuring those results: What is truly offloaded?

Climbing the autonomy ladder: The **quantity** of security tasks **driven and completed** by the AI instead of a human becomes progressively greater



AI Metrics that Matter: A pyramid for measuring value

Security Operations

- Median Time to Detect
- Median Time to (first) Respond
- Median Time to Contain

Cloud Security

- Day 0 Time To Value
- Time To Remediation
- Asset Coverage
- Attack Exposure

Applying agentic AI from here

- Next week you should:
  - Define your Metrics that Matter (e.g., reduce time to detect)
- In the first three months following this presentation you should:
  - Write down the steps you take to achieve that outcome today
- Within six months you should:
  - Experiment with any agent-builder, pointing it at those steps

RSAC 2025 Conference

Many Voices.
One Community.

Thank you